A simple mathematical description of epidemics using adapted Bateman equations which describe decaying cascades of nuclear excited states

P.Petkov
"Horia Hulubei" National Institute for Physics and Nuclear Engineering,
R-76900, Bucharest-Măgurele, Romania

Abstract

A mathematical formalism is presented which allows studies in multistep processes where one particular group of objects is generated and in turn generates a new group with similar properties with some rates. Such multistep processes are observed for example in epidemiology. For this purpose an adaptation is made of the Bateman equations describing a system of nuclear levels interconnected by transitions. Closed mathematical expressions allow an easy calculation/prediction of the properties of such multistep processes.

The need of adequate mathematical description/modeling of epidemics evolution appeals strongly for developments in this field, it is enough mentioning the pandemic COVER-19 in the beginning of 2020 (see e.g. the report[1] of the WHO.) The author of the present work is an amateur in epidemiology, but a brief consideration of recent papers in that field employ a variety of mathematical apparatus to analyze the statistics of the available data in order to provide explanations for it and predict possible future scenarios (cf. e.g. Refs. [2, 3, 4, 5] and references therein). While an analysis of the initial part of the experimental curve describing the number of observed cases as function of time is relatively reliable for short range forecasts, also when the forecast is very close to its maximum (peak), in general it is desirable to dispose with description of the epidemics in the full time range till its disappearing (falling in the background). More involved epidemics models [6] involving also solving of non-linear differential equations (e.g. SIR) are given in Refs. [7, 8].

One different possibility for epidemics description is provided by adapting the Bateman equations known since the beginning of the 20th century (see e.g.[9], also [10] for more details), which are linear, for the epidemiology case. The Bateman equations describe the time evolution of the population of a system of nuclear levels which are interconnected by transitions (e.g. emission of gamma-rays in Nuclear physics, also photons in similar systems in Atomic physics). Namely, the system of N nuclear levels is governed by N differential equations (i=1...N):

$$\dot{n}_i(t) = -\lambda_i n_i(t) + \sum_h \lambda_h b_{hi} n_h(t) \tag{1}$$

where $n_i(t)$ is the evolution of the population of the level of interest i, the levels h ($h > i$) are direct feeders of that level, $\lambda_{i,h}$ are the decay constants and b_{hi} are the branching ratios. The general solution (if $\lambda_k \neq \lambda_i$ for $i \neq k$) including all possible feeding levels can be written in a straightforward way (c.f. e.g. ref.[9]) as

$$n_i(t) = \sum_{k \geq i} C_{ik} e^{-\lambda_k t} = (n_i(0) - \sum_{k>i} C_{ik}) e^{-\lambda_i t} + \sum_{k>i} C_{ik} e^{-\lambda_k t} \tag{2}$$

where $n_i(0)$ is the initial population (at $t = 0$) of the level i, the numbering of the levels starts from it and the coefficients C_{ik} are given by the recursive equations ($k > i$)

$$C_{ik} = \frac{\lambda_k}{\lambda_i - \lambda_k} \left(b_{ki} n_k(0) - \sum_{m>k} C_{km} b_{ki} + \sum_{l>i}^{k-1} b_{li} \frac{\lambda_l}{\lambda_k} C_{lk} \right) \tag{3}$$

On the other hand, provided the functions $n_h(t)$ of the direct feeders h are known the solution of the set of equations (1) for every particular level i could be written as:

$$n_i(t) = n_i(0) e^{-\lambda_i t} + \sum_h \lambda_h b_{hi} \int_0^t dx n_h(x) e^{-\lambda_i (t-x)} \tag{4}$$

In the latter formula, no details about the history of the population of the feeding levels h are present. However, this information is contained in the functions $n_h(t)$ and that is the way the function $n_i(t)$ "feels" the earlier stages of the cascade history.

We use the above approach to describe in very crude way a process which has the features of epidemics. Only two decay rates λ_R (rate of recovery) and λ_C (rate of contamination) are employed which are assumed to be constant in time. Of course, the latter may be not true in the history of each case of disease when going from the begin of the illness, the peak (crisis) and the recovery (see discussion below). Let us consider the evolution of the different generations of contaminated people with time. At time t=0, there are $n_1(0)$ cases which recover with a constant rate λ_R. In practice, such assumption may be very crude since: (i) the time necessary for the recovery process depends on the health parameters of the patient in the more general sense. (ii) the rate of recovery may be also not a constant i.e. the illness goes to different stages during the recovery, characterized by different health parameters (iii) application of treatments with medicaments with different efficiencies. Nevertheless, an average of λ_R's can be used in calculations as done below. Also, some upper limit λ_R fixes the time limits of the full process of acquiring immunity for the whole population. The same considerations apply to the rate of contamination λ_C which, however, may be very well influenced by measures as confinement of ill people, reduction of social contacts etc. With these remarks the time evolution of $n_1(t)$ is governed by the well known radioactivity decay law of Rutherford-Soddy:

$$\frac{dn_1(t)}{dt} = -\lambda_R n_1(t) \tag{5}$$

with the solution $n_1(t) = n_1(0)e^{-\lambda_R t}$. Here, the mutual new contamination of already contaminated people is not taken into account, of course.

The population of the second generation with time $n_2(t)$ is governed by the equation

$$\frac{dn_2(t)}{dt} = -\lambda_R n_2(t) + \lambda_C n_1(0)e^{-\lambda_R t} \tag{6}$$

It is easy to show that its solution, with the boundary condition $n_2(0)$=0 is given by:

$$n_2(t) = \lambda_C n_1(0) t e^{-\lambda_R t} \tag{7}$$

The third generation population $n_3(t)$ is governed by the equation

$$\frac{dn_3(t)}{dt} = -\lambda_R n_3(t) + \lambda_C n_2(t) \tag{8}$$

Again it is easy to show that its solution is given by:

$$n_3(t) = \frac{{\lambda_C}^2 n_1(0) t^2 e^{-\lambda_R t}}{2}, \tag{9}$$

for the fourth generation one obtains $n_4(t) = {\lambda_C}^3 n_1(0) t^3 e^{-\lambda_R t}/(2.3)$ and so on. The final result for the n^{th} generation obtained by induction reads:

$$n_n(t) = \frac{{\lambda_C}^{n-1} n_1(0) t^{n-1} e^{-\lambda_R t}}{(n-1)!}. \tag{10}$$

The summing up the number of contaminated people from different generations one obtains for their total number as function of time:

$$N(t) = \sum_{k=1,2,..N_{Max}} \frac{(\frac{\lambda_C}{\lambda_R})^{k-1} n_1(0) x^{k-1} e^{-x}}{(k-1)!}. \tag{11}$$

where N_{Max} is the last generation and $x = \lambda_R t$. This expression represents a sum of Poisson distributions weighted by the factors $(\frac{\lambda_C}{\lambda_R})^{k-1}$.

Table 1: Calculations for contamination rate $\lambda_C = 20$ contaminations per day caused by an ill person and different mean recovery times τ_R ($\lambda_R = 1/\tau_R$). The initial population $n_1(0) = 1$. In the second column, the the contaminated generation number n is shown followed by the position of the peak maximum in time t_{Max}. For the other quantities see text. The table is left-aligned and no abbreviated formats for the large numbers (as e.g. a*10^b) are used in order to get better optical effect on the observed increase with n and τ_R.

τ_R	n	t_{Max} [days]	$(n-1)^{n-2}e^{-(n-1)}/(n-2)!$	$(\lambda_C/\lambda_R)^{(n-1)}$	n_n^{Max}
7 days	2	7.0	0.368	141	52
	3	14.1	0.135	19837	2684
	4	21.1	0.075	2793990	208656
	5	28.2	0.049	393519814	19220178
	6	35.2	0.035	55425326023	1945067235
14 days	2	14.1	0.368	281	103
	3	28.2	0.135	79349	10738
	4	42.3	0.075	22351925	1669255
	5	56.3	0.049	6296317036	307522851
	6	70.4	0.035	1773610432745	62242151525
21 days	2	20.8	0.368	417	153
	3	41.7	0.135	173611	23496
	4	62.5	0.075	72337962	5402242
	5	83.3	0.049	30140817901	1472128897
	6	104.2	0.035	12558674125514	440727503315

It can be shown that the function $n_n(t)$ has one maximum which occurs at

$$t_{max} = (n-1)/\lambda_R = (n-1)\tau_R \tag{12}$$

where with τ_R we denote the reciprocal of λ_R and which is equal to the mean time needed for recovery (analog of the mean lifetime of an excited nuclear state). Thus, at $\tau_R = 14$ days the 4^{th} generation of contaminated people will have its maximum (peak) 42 days after the begin of the contamination, the 5^{th} one - after 70 days and so on within our simple model. The height of the peak maximum at t_{max} is given by

$$n_n(t) = (\frac{\lambda_C}{\lambda_R})^{n-1} n_1(0) \frac{(n-1)^{n-2}}{(n-2)!} e^{-(n-1)}. \tag{13}$$

and a more careful consideration (see also Table 1) reveals that its value depends extremely strong on the ratio (λ_C/λ_R) at the power of $(n-1)$. For example, if that ratio is 1000, for the 4^{th} generation the result will lead to 10^9. The factor $(n-1)^{n-2}e^{-(n-1)}/(n-2)!$ will reduce that number only by about 10%. Another interesting feature of the formalism is that in the composition of $N(t)$, the total number of contaminated people, the last generation of contaminated $n_{N_max}(t)$ dominates also strongly the other terms of the sum (with a factor of 100 and more, Table 1) at any time. Then, the peak position of $n_{N_max}(t)$ dominates the peak position of the development in time of the whole epidemics, indeed. Some more details, e.g. on the time range of the epidemics, are presented in Fig. 1

This paper does not intend to investigate and provide predictions for the development of concrete epidemics and pandemics. Such tasks are beyond its scope and the competence of the author. But our results may be used by specialists in epidemiology when the remarks made above about the rates λ_R, λ_C are taken into account. Moreover, one has to consider important factors as specificities in every country coexisting with globalization effects, the existence of different social groups and different health systems in each country etc. From that point of view, we think that the mathematical formalism proposed here can help the specialists in epidemiology to solve problems and if necessary, they can introduce modifications in the approach to take into account more complex situations.

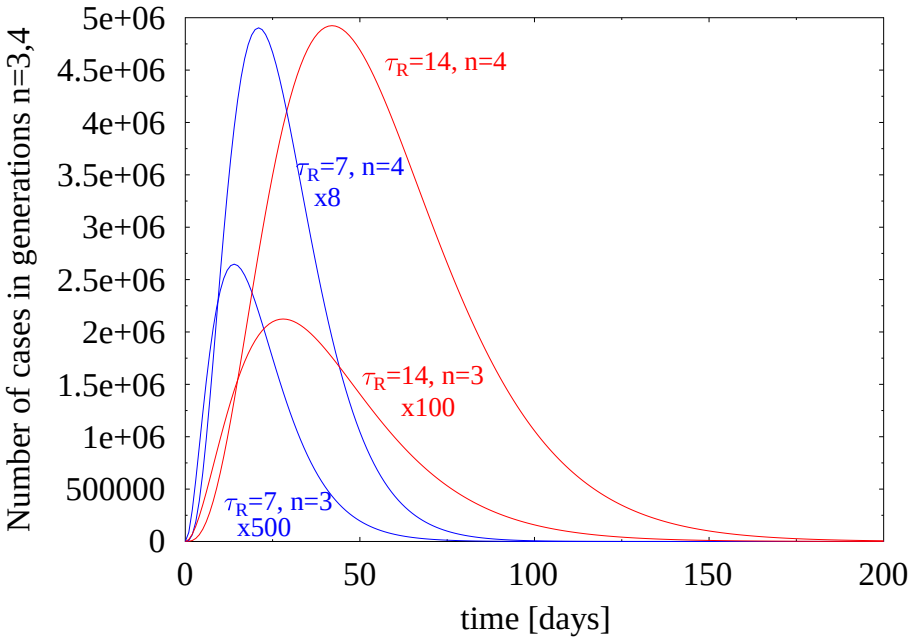

Figure 1: A graphical representation of $n_n(t)$ functions of third an fourth generation combined in pairs with common τ_R (7 or 14 days) and a common λ_C=20 for all four cases. Note the different factors used to scale apart that of τ_R=14, n=4. The time extent of the epidemics for τ_R=7 and τ_R=14 for the cases displayed is about 100 and 200 days, respectively. See also text.

References

[1] WHO, Corona virus disease (COVID-19). 2020. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200226-sitrep-37-covid-19.pdf?sfvrsn=2146841e-2. Accessed Date: 20 March, 2020.

[2] W.W.Koczkodaj et al, Global Epidemiology, in Press. //doi.org/10.1016/j.gloepi.2020.100023

[3] Heymann, D.L, Shindo, N. COVID-19: What is next for public health? The Lancet. 2020;395:542-545. https://doi.org/10.1016/S0140-6736(20)30374-3.

[4] Murthy, S, Gomersall C.D, Fowler, R.A. Care for Critically Ill Patients with COVID-19. JAMA. 2020; In press. https://doi:10.1001/jama.2020.3633.

[5] Kakiashvili, T, Koczkodaj, W.W, Woodbury-Smith, M. Improving the Medical Scale Predictability by the Pairwise Comparisons Method: Evidence from a Clinical Data Study. Computer Methods and Programs in Biomedicine. 2012;105:210-216.

[6] Kermack, W. O. McKendrick, A. G. Proceedings of the Royal Society A 115 (1927) 700

[7] Hethcote H , SIAM Review. 42 (4) (2000) 599

[8] Harco Tiberiu; Lobo, Francisco S. N.; Mak, M.K., Applied Mathematics and Computation 236 (2004) 184

[9] E. Segre, in Nuclei and Particles, W.A.Benjamin, Inc., New York 1965, p.157

[10] P.Petkov et al., Nuclear Instruments and Methods 321 (1992) 259